

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

in search, using sw model

cember 30, 2003, 12:08:39 ; Search time 21 Seconds
(without alignments)
311.403 Million cell updates/sec

-09-933-638a-11

2 LEGVLTHQOFSSYEPELFP.....RAEYIEAFENIYILKGRFK 68

OSUM62

pop 10.0 , Gapext 0.5

3308 seqs, 96168682 residues

ts satisfying chosen parameters: 283308

3th: 0

3th: 2000000000

imum Match 0%

aximum Match 100%

isting first 45 summaries

IR 76:*

pir1:*

pir2:*

pir3:*

pir4:*

the number of results predicted by chance to have a
r than or equal to the score of the result being printed,
ad by analysis of the total score distribution.

SUMMARIES

ary	ch	Length	DB	ID	Description
0.0	297	1	151648		transcription init
0.0	302	1	JC5513		transcription init
0.0	316	1	S34437		transcription init
0.0	339	1	TWU02D		transcription init
3.8	302	1	JC4059		transcription init
5.7	353	1	A35615		transcription init
2.5	340	2	A48671		transcription init
1.9	231	2	A35873		transcription init
0.7	240	2	A30366		transcription init
1.7	258	2	JQ1666		transcription init
3.2	200	2	S61088		transcription init
3.2	201	2	S30216		transcription init
3.9	200	2	S21140		transcription init
3.9	215	2	T03386		transcription init
3.9	233	2	S23522		transcription init
1.1	200	2	S10945		transcription init
7.8	200	1	TW020D		transcription init
7.5	200	2	S10946		transcription init
3.3	249	2	E90102		TATA binding prote
3.3	249	2	E90137		TATA binding prote
3.3	249	2	G90118		transcription fact
3.7	224	2	S37740		transcription init
3.8	246	2	S41473		transcription init
3.8	233	2	S52407		TATA box-binding p
3.5	222	2	F72572		probable TATA-box
3.8	228	2	JQ2124		transcription init
1.3	181	2	E90246		TATA box binding p
1.3	198	2	S55311		TATA-binding prote
3.9	191	2	D71093		probable TATA-bind

30	136	39.8	183	2	C64363	TATA-bindin
31	134.5	39.3	191	2	E75072	transcript:
32	134	39.2	183	2	E69296	transcript:
33	133	38.9	527	2	T22000	hypothetic
34	132	38.6	190	2	JC4514	TATA-bindin
35	132	38.6	191	2	A54275	TATA-bindin
36	132	38.6	191	2	D47230	transcript:
37	130.5	38.2	181	2	D69084	TATA-bindin
38	125	36.5	186	2	A84375	transcript:
39	121	35.4	100	2	T08248	TATA-bindin
40	121	35.4	186	2	T08331	TATA-bindin
41	120	35.1	186	2	T08255	TATA-bindin
42	120	35.1	186	2	T44922	TATA box b:
43	116.5	34.1	181	2	T08317	TATA-bindin
44	110.5	32.3	186	2	JG0162	TBP-like p
45	84.5	24.7	182	2	T31320	TATA-bindin

ALIGNMENTS

RESULT 1

151648
transcription initiation factor IID - African clawed frog
N;Alternate names: TATA-binding protein
C;Species: Xenopus laevis (African clawed frog)
C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 03-Mar
C;Accession: I51648; S68859; S26292
R;Hashimoto, S.; Fujita, H.; Hasegawa, S.; Roeder, R.G.; Horikoshi, M.
Nucleic Acids Res. 20, 3788, 1992
A;Title: Conserved structural motifs within the N-terminal domain of T
A;Reference number: I51648; MUID:92350691; PMID:1641350
A;Accession: I51648
A;Status: translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-297 <HAS>
A;Cross-references: EMBL:X66033; NID:965148; PIDN:CAA46832.1; PID:9651
A;Note: only a part of the translation is shown
R;Labhart, P.
FEBS Lett. 386, 110-114, 1996
A;Title: Phosphorylation of the N-terminal domain of Xenopus TATA-box
A;Reference number: S68859; MUID:96228045; PMID:8647263
A;Accession: S68859
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-54,'R',56-176,'I',178-297 <LAB>
A;Note: only a list of differences from sequence S26292 is shown in th
C;Genetics:
C;Gene: TFIIDtau
C;Superfamily: human transcription initiation factor IID
C;Keywords: DNA binding; nucleus; transcription initiation

Query Match 100.0%; Score 342; DB 1; Length 297;
Best Local Similarity 100.0%; Pred. NO. 1.7e-32;
Matches 68; Conservative 0; Mismatches 0; Indels 0; G

QY	1	LEGVLTHQOFSSYEPELFPGLYRMKPRIVLLIFVSGKVLTKAKVRAIIEAF
Db	228	LEGVLTHQOFSSYEPELFPGLYRMKPRIVLLIFVSGKVLTKAKVRAIIEAF
QY	61	PILKGRK 68
Db	288	PILKGRK 295

RESULT 2

JC5513
transcription initiation factor IID - chicken
N;Alternate names: TATA-binding protein
C;Species: Gallus gallus (chicken)
C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 16-Jun
C;Accession: JC5513
R;Yamauchi, J.; Sugita, A.; Fujiwara, M.; Suzuki, K.; Matsumoto, H.; Y

C:Species: Homo sapiens (man)
C:Date: 20-Jul-1990 #sequence revision 19-May-1995 #text_change 18-Feb
C:Accession: A34830; A34831; S10944; I60128
R:Peterson, M.G.; Tanese, N.; Pugh, B.F.; Tjian, R.
Science 248, 1625-1630, 1990

A;Title: Functional domains and upstream activation properties of clo
A;Reference number: A34830; MUID:90302006; PMID:2363050

A:Accession: 1-359 (FEI)
A:Cross-references: GB:M5654; NID:g339491; PIDN:AAA36731.1; PID:g339491; R:Kao, C.C.; Lieberman, P.M.; Schmidt, M.C.; Zhou, Q.; Pei, R.; Berk, S.; Title, Cloning of a transcriptionally active human TATA binding factor. Science 248, 1646-1649, 1990
A:Reference number: A34831; MUID:90302010; PMID:21942899

A;Title: Cloning of a transcriptionally active human TATA binding factor cDNA.
A;Reference number: A34831; MUID:90302010; PMID:2194289
A;Accession: A34831

A; Molecule type: DNA
A; Residues: 1-17, 'N', 19-186, 'R', 188-339 <KAO>
R; Hoffmann, A.; Sinn, E.; Yamamoto, T.; Wang, J.; Roy, A.; Horikoshi,

A;Title: Highly conserved core domain and unique N terminus with pres
A;Reference number: S10944; MUID:90326195; PMID:2374612
A;Accession: S10944

A;Accession: S10944
A;Molecule type: mRNA
A;Residues: 1-91.96-339 <HOP>
A;Reference number: S10944; MUID:90326195; PMID:2374612

Science 248, 1626, 1990
A; Title: Cloning of the human TATA binding factor: Expression of a tr
A; Reference number: 160128

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-186, 'R', 188-299, 'MIKPR', 300-339 <RES>

A;Gene: GDB:TBP; GTF2D1
A;Cross-references: GDB:138768; OMIM:600075
A;Map position: 6q27-6q27

Query Match 100.0%; Score 342; DB 1; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-32;

Best Local Similarity 100.0%; Pred. No. 2e-32;
Matches 68; Conservative 0; Mismatches 0; Indels 0;

Qy 61 PILGFRK 68
||| |||
Db 330 PILGFRK 337

RESULT 5
JC4059

N;Alternate names: TATA-binding protein
C;Species: *Trimeresurus gramineus* (Indian green tree viper)
C;Date: 18-Feb-2000 #sequence revision 18-Feb-2000 #text change 16-Jun-2000
transcription initiation factor IID = Indian green tree viper

R; Nakashima, K.; Nobuhisa, I.; Deshimaru, M.; Ogawa, T.; Shimohigashi, T.; Akita, Y.; Ohtsuka, S.; Nishimura, H.; Nakamura, A.; Nakamura, T.
Gene 152, 209-213, 1995

Gene 152, 209-213, 1995
A;Title: Structures of genes encoding TATA box-binding proteins from :
A;Reference number: JC4059; MUID:95137390; PMID:7835702
A;Accession: JC4059

A;Residues: 1-302 <NAK>
A;Cross-references: DDBJ:D31776; NID:G1483196; PIDN:BAA06554.1; PID:98
C;Comment: This protein recognizes and binds the TATA box and is requi

C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 21-Ju
 C:Accession: A48671; T16902
 R.Lichtsteiner, S.; Tjian, R.
 Proc. Natl. Acad. Sci. U.S.A. 90, 9673-9677, 1993
 A:Title: Cloning and properties of the Caenorhabditis elegans TATA-bo
 A:Reference number: A48671; MUID:94022438; PMID:8415761
 A:Accession: A48671
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-340 <LIC>
 A:Cross-references: GB:L07754; NID:g156447; PIDN:AAA03582.1; PID:g156
 R:Du, Z.
 submitted to the EMBL Data Library, June 1994
 A:Description: The sequence of C. elegans cosmid T20B12.
 A:Reference number: S46772
 A:Accession: T16902
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-340 <DUZ>
 A:Cross-references: EMBL:U10401; NID:g500713; PID:g500715; PIDN:AAA19
 A:Experimental source: strain Bristol N2
 C:Genetics:
 A:Introns: 70/1; 257/3; 312/3
 C:Superfamily: transcription initiation factor IID
 C:Keywords: DNA binding; nucleus; transcription initiation

Query Match 82.5%; Score 282; DB 2; Length 340;
 Best Local Similarity 77.9%; Pred. No. 2.1e-25;
 Matches 53; Conservative 8; Mismatches 7; Indels 0;

QY 1 LEGVLVTHQFSSYPPELPGLIYRMKPRIVLLIFVSGKVLTGAKVRAEIV
 DB 273 LEGLCITHSQSFSTYEPPELPGLIYRMKPRIVLLIFVSGKVLTGAKVRAEIV
 QY 61 PILKGFRK 68
 DB 333 PILKGFKK 340

RESULT 8
 A35873
 transcription initiation factor IID [similarity] - fission yeast (Sch
 N:Alternate names: TFIID protein
 C:Species: Schizosaccharomyces pombe
 C:Date: 09-Nov-1990 #sequence_revision 13-Jan-1993 #text_change 28-Ju
 A:Accession: A35873; S10940; T38509; S10218; S62514
 R:Hofmann, A.; Horikoshi, M.; Wang, C.K.; Schroeder, S.; Weil, P.A.;
 Genes Dev. 4, 1141-1148, 1990
 A:Title: Cloning of the Schizosaccharomyces pombe TFIID gene reveals
 A:Reference number: A35873; MUID:91007258; PMID:2210373
 A:Accession: A35873
 A:Molecule type: DNA
 A:Residues: 1-231 <HOF>
 A:Cross-references: GB:X53415; NID:g5114; PIDN:CAA37494.1; PID:g29593;
 A:Note: the authors translated the codon TCC for residue 123 as Lys, i
 R:Fikes, J.D.; Becker, D.M.; Winston, F.; Guarente, L.
 Nature 346, 291-294, 1990
 A:Title: Striking conservation of TFIID in Schizosaccharomyces pombe
 A:Reference number: S10940; MUID:90326164; PMID:2197558
 A:Accession: S10940
 A:Molecule type: mRNA
 A:Residues: 1-231 <NAT>
 R:Jones, L.; Murphy, L.; McNeil, A.; Simpson, I.; Harris, D.; Barrell,
 submitted to the EMBL Data Library, October 1995
 A:Reference number: Z21798
 A:Accession: T38509
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-231 <J02>
 A:Cross-references: EMBL:Z66525; PIDN:CAA91430.1; PID:g1044934; GSPDB
 A:Experimental source: strain 972h(-); cosmid c29E6
 R:McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Murphy,
 submitted to the EMBL Data Library, January 2000

an transcription initiation factor IID
 nding; nucleus; transcription initiation
 98.8%; Score 338; DB 1; Length 302;
 arity 98.5%; Pred. No. 5.2e-32;
 onservative 0; Mismatches 1; Indels 0; Gaps 0;
 VLT HQFSSYPPELPGLIYRMKPRIVLLIFVSGKVLTGAKVRAEIVAFENIY 60
 VLT HQFSSYPPELPGLIYRMKPRIVLLIFVSGKVLTGAKVRAEIVAFENIY 292
 GFRK 68
 GFRK 300

iation factor IID - fruit fly (Drosophila melanogaster)
 TATA box-binding protein
 0 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000
 5; A38416; J04056
 ht, B.D.; Peterson, M.G.; Pugh, B.P.; Tjian, R.
 , 1990
 and characterization of the Drosophila gene encoding the TATA box bin
 : A35615; MUID:90304877; PMID:2194666

RNA
 <HOE>
 : GB:M38082; NID:g158541; PIDN:AAA28931.1; PID:g158542
 da, C.T.; Horikoshi, M.; Roeder, R.G.; Parker, C.S.
 Sci. U.S.A. 87, 9148-9152, 1990
 e encoding Drosophila transcription factor TFIID.
 : A38416; MUID:91067664; PMID:2123550

RNA
 <MUH>
 : GB:M38388; NID:g158531; PIDN:AAA28926.1; PID:g158532
 l.; Burke, T.W.; Kadonaga, J.T.
 1995
 of the genes encoding transcription factor IIB and TATA box-binding p
 : J04055; MUID:95180719; PMID:7875589

NA
 <LIR>
 : GU:U11718; NID:g515664; PIDN:AAA68629.1; PID:g515665

P
 : FlyBae:FBgn0031687
 sophila transcription initiation factor IID
 nding; nucleus; transcription initiation
 85.7%; Score 293; DB 1; Length 353;
 arity 83.8%; Pred. No. 1.1e-26;
 onservative 6; Mismatches 5; Indels 0; Gaps 0;
 VLT HQFSSYPPELPGLIYRMKPRIVLLIFVSGKVLTGAKVRAEIVAFENIY 60
 VLT HQFSSYPPELPGLIYRMKPRIVLLIFVSGKVLTGAKVRAEIVAFENIY 343
 GFRK 68
 KFKK 351

iation factor IID - Caenorhabditis elegans
 TATA-binding protein
 abditis elegans


```

transcription factor, TATA-binding
s (maize)
6 #sequence_revision 23-Feb-1996 #text_change 26-Aug-1999
8; S32622
x, G.
4-298, 1992
rent cDNAs encoding TFIIID proteins of maize.
: S21140; MUID:92249585; PMID:1577169
8
RNA
<HAA>
h, B.; Cigan, M.; Freeling, M.
MBL Data Library, April 1993
two genes encoding maize TATA-binding protein, expressed differential
: S32622
2
RNA
<VOG>
GB:113302; EMBL:Z22172; NID:q293905; PIDN:AAA65942.1; PID:q293906
transcription initiation factor IID
binding, transcription initiation
79.2%; Score 271; DB 2; Length 200;
arity 77.9%; Pred.No.2.4e-24;
conservative 5; Mismatches 10; Indels 0; Gaps 0;
VLTHTQOFSSEPELPGLIYMKERTVLLIFVSGKVLTCAKVRAEITAFENIY 60
AYSHGAFSSSEPELPFGIYIRKMPKIVLLIFVSGKVLTCAGVREETVAFENIY 189
GFRK 68
|||
EPRK 197

```

```

Initiation factor IID - wheat  
DNA-binding protein TFIID  
in aestivum (common wheat)  
#sequence_revision 20-Feb-1995 #text_change 15-Oct-1999  
  
Chase, J.A.; Chase, M.R.; Davis, E.A.; Ackerman, S.  
J Biol Chem 271:1494, 1993  
IID TATA binding protein.  
 : S30216; MUID:93219135; PMID:8464747  
 :  
ary  
RNA  
APS>  
EMBL:Z18804; NID:g21876; PIDN.CAA79268.1; PID.g21877  
scription initiation factor IID  
nding; nucleus; transcription initiation  
  
79.2%; Score 271; DB 2; Length 201;  
arity 77.9%; Pred.No.2.4e-24;  
nservative 5; Mismatches 10; Indels 0; Gaps  
  
LTHQOFSYPELPGLIYRMKPRIVLLIFVSGKWLTGAKVRAEITYEAFFNY  
:  
AYSHGAFSYPPELPGLIYMRQPVLIIFVSXKI VLGAKVRETYSAFNYY  
:  
GRFK 68  
|||  
EFKK 198
```

transcription factor, TATA-binding
 factor IID.1 - maize
 (maize)
 #sequence revision 23-Feb-1996 #text change 26-Aug-1999

C;Accession: S21140; S32639
R;Haass, M.M.; Feix, G.
FEBS Lett. 301, 294-298, 1992
A;Title: Two different cDNAs encoding TFIID proteins of maize.
A;Reference number: S21140; MUID:92249585; PMID:1577169
A;Accession: S21140
A;Molecule type: mRNA
A;Residues: 1-200 <HAA>
R;Vogel, J.M.; Roth, B.; Cigan, M.; Freeling, M.
submitted to the EMBL Data Library, April 1993
A;Description: The two genes encoding maize TATA-binding protein, exp:
A;Reference number: S32622
A;Accession: S32639
A;Molecule type: mRNA
A;Residues: 1-200 <VOG>
A;Cross-references: GB:I13301; EMBL:Z22149; NTD:Q293903; PIDN:AAA6594;
C;Superfamily: transcription initiation factor IID
C;Keywords: DNA binding; transcription initiation

Query Match 78.9%; Score 270; DB 2; Length 200;
Best Local Similarity 77.9%; Pred. No. 3.2e-24;
Matches 53; Conservative 5; Mismatches 10; Indels 0;

QY 1 LEGVLVTHQFSSVEPELPGLIYRMKPIRVLLIFVSGKVLTCAKVRAEIYEAI
 |||||?|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 130 LEGLAYSHGHAFSSVEPELPGLIYRMKQPKIVLLIFVSGKIVLTGAQRRETYTAL

QY 61 PILKGFRX 68
Db 190 PVLAEFRK 197
 :|||
 :|||

```

RESULT 14
T03386
transcription initiation factor tbpl - maize
N:Alternate names: TATA box-binding protein
C:Species: Zea mays (maize)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 26-Aug-
C:Accession: T03386
C:Author: R:Godde-meier, M.L.; Feix, G.
Gene 174, 111-114, 1996
A:Title: Genomic structure of the maize TATA-box binding protein 1 (TE
A:Reference number: Z14917; MUID:97017135; PMID:8863736
A:Accession: T03386
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-215 <GOD>
A:Cross-references: EMBL:X30652; NID:g945021; PIDN:CAA62224.1; PID:g97
A:Experimental source: cultivar rtc2; isolate DK105; endosperm
C:Genetics:
A:Gene: tbpl
A:Introns: 25/2; 54/3; 78/3; 95/3; 110/3; 141/3; 175/3; 215/2
C:Superfamily: transcription initiation factor IID
C:Keywords: DNA binding; transcription initiation

```

```

Query Match          78.9%;   Score 270;   DB 2;   Length 215;
Best Local Similarity 77.9%;   Pred. No. 3.4e-24;
Matches 53;   Conservative 5;   Mismatches 10;   Indels 0;   G

QY 1  LEGVLVTHQOFSFSEYPELPGFIYRMKIPRVLIIIFVSGKVLTGAKVRAIEVEAF
    |||||
DB 130 LEGLAYSHGAFSEYPELPGFIYRMKQPKVLIIFVSGKIVLTGAKVREETVAF
    |||||

QY 61 PILKGFRK 68
    |||||
DB 190 PVLAERFK 197

```

RESULT 15
S23522
transcription initiation factor IID - wheat
N:Alternate names: TATA-binding protein TFIID
C:Species: Triticum aestivum (common wheat)

94 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
22; S16759
amd, M.; Tamura, T.; Sumita, K.; Iwabuchi, M.
19, 867-872, 1992
n and characterization of a cDNA clone encoding the TATA box-binding pr
r: S23522; MUID:92353396; PMID:1643287
22
mRNA
<KAW>
s: EMBL:X59874; NID:g21874; PIDN:CAA42531.1; PID:g21875
anscription initiation factor IID
inding; nucleus; transcription initiation
larity 78.9%; Score 270; DB 2; Length 233;
Conservative 7; Mismatches 9; Indels 0; Gaps 0;
LVLTHQOFSSYEPELFPGLIYRMKPRIVLIFVSGKVLGTGAKVRAEIVAEFENIY 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
LAYSHGAFSSYEPELFPGLIYRMKQPKIVLIVFVSGKIVLTGAKVRDEIYAAEFENIY 222
KGPRK 68
:||
TEYRK 230
December 30, 2003, 12:11:47
s